INSTITUTE OF PHYSICS PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **47** (2002) 3503–3517 PII: S0031-9155(02)38653-6

Segmentation of MR images for computer-assisted surgery of the lumbar spine

C L Hoad and A L Martel

Department of Medical Physics, University Hospital, Queen's Medical Centre, Nottingham, UK

Received 25 June 2002
Published 18 September 2002
Online at stacks.iop.org/PMB/47/3503

Abstract

This paper describes a segmentation algorithm designed to separate bone from soft tissue in magnetic resonance (MR) images developed for computer-assisted surgery of the spine. The algorithm was applied to MR images of the spine of healthy volunteers. Registration experiments were carried out on a physical model of a spine generated from computed tomography (CT) data of a surgical patient. Segmented CT, manually segmented MR and MR images segmented using the developed algorithm were compared. The algorithm performed well at segmenting bone from soft tissue on images taken of healthy volunteers. Registration experiments showed similar results between the CT and MR data. The MR data, which were manually segmented, performed worse on visual verification experiments than both the CT and semi-automatic segmented data. The algorithm developed performs well at segmenting bone from soft tissue in MR images of the spine as measured using registration experiments.

1. Introduction

The use of computer-assisted surgery (CAS) for pedicle screw insertion in the lumbar spine has improved screw placement *in vivo* (Laine *et al* 1997, Schwarzenbach *et al* 1997, Merloz *et al* 1998, Amiot *et al* 2000). By registering pre-operative computed tomography (CT) images to the patient's frame of reference in the operating theatre, positions and projections of instruments trajectories can be seen in real time during the operation. A move away from CT to magnetic resonance (MR) for diagnostic imaging of spinal disorders has prompted development of MR images, which can be used with CAS (Martel *et al* 1998, Hoad *et al* 2000, 2001). These images have been shown to be geometrically accurate (Martel *et al* 1998, Hoad *et al* 2000) and produce good tissue contrast between the bone and surrounding soft tissue of the spine (Hoad *et al* 2001). These are two important qualities needed in the images for CAS to be successful.

Many of the systems developed for CAS of the spine use surface registration to obtain the transformation between the image and patient's frame of reference (Amiot *et al* 1995,

Nolte *et al* 1995, Merloz *et al* 1997). This produces reliable registration and is not dependent on how well the surgeon can identify a set of pre-defined anatomical landmarks as in paired-point matching. A three-dimensional (3D) model of the bone surface must be generated from the pre-operative images for surface registration to be used. Digitized surface points from physical space are registered to the computer generated surface from the images. Using CT images, this surface can be generated by selecting an image intensity threshold level, which separates the signal of bone from that of soft tissue. For MR images the problem is more difficult. Heterogeneous signal intensities in tissues, poor signal to noise ratios (SNR) and non-uniform signal intensities make the problem of bone segmentation complex. Although manual segmentation of the MR data using a 'live wire' technique (Barrett *et al* 1997) has been carried out successfully (Hoad *et al* 2001), the technique is not suitable for CAS as the process is extremely time consuming and user dependent.

Segmenting MR images into different tissue types has been studied extensively (Bezdek *et al* 1993, Clarke *et al* 1993, 1995, Manousakas *et al* 1998, Martel *et al* 1999). Much of this work has concentrated on segmenting brain tissue into its various components, grey matter, white matter, csf etc (Brandt *et al* 1994, Lemieux *et al* 1999, Calmon *et al* 2000). The use of more than one dataset with different tissue contrast has made multi-spectral approaches to this problem an area of intense research (Vannier *et al* 1985, Choi *et al* 1991, Hall *et al* 1992, Liang 1993, Mitchell *et al* 1994). The segmentation problem for MR CAS images is a single contrast one. Only one dataset is available to manipulate. Techniques for this kind of segmentation include edge detection, boundary tracing, thresholding, seed growing, template models and morphological filtering (Sonka *et al* 1993, Clarke *et al* 1995). These techniques applied individually have many limitations (Clarke *et al* 1995), however when optimized for specific parts of the segmentation process and combined into a single algorithm, the results can be excellent (Kapur *et al* 1996, Lemieux *et al* 1999). This type of approach is needed for the problem of segmenting the bone from soft tissue in CAS MR images of the spine, as the image contrast and quality do not allow for a single technique, other than manual segmentation, to be used successfully.

Many different approaches have been taken to the assessment of segmentation accuracy. The traditional approach is to compare the segmentation carried out using a new technique with a gold standard, typically manual segmentation carried out by experts. This approach is unsatisfactory in this study for two reasons. Manual segmentation of the MR images is very subjective making it a poor gold standard and a comparison of two segmented images on a pixel by pixel basis is unlikely to provide information about whether the segmented images are fit for their intended purpose, i.e. surface registration. The standard approach to assessing the accuracy of surface-based registration between physical space and image space using CAS systems is to implant fiducial markers into a physical phantom, which is then scanned using CT (Herring *et al* 1998, Bachler *et al* 2001). Surface points on the phantom are digitized using the CAS system and the bone surface is extracted from the CT data using simple thresholding. Surface-based registration is used to generate a coordinate transform function (CTF) that minimizes the distance between the digitized surface points and the bone surface generated from the image data. The position of the fiducial markers in physical space is determined using a tracking device and the position of the markers in the CT image is calculated after transformation using the CTF. The difference between the positions after registration is used to calculate the registration accuracy. This approach works well for CT images as the properties of the images obtained from a plastic phantom are not significantly different to those of patient images. In MRI this is not the case as there is a complex relationship between tissue type and image contrast, and even a cadaver spine has a very different appearance to a spine *in vivo* (Martel *et al* 1998). In order to overcome this problem, we have scanned a patient first and

then constructed a phantom from the image data, rather than constructing the phantom first and then scanning it.

The aim of the work presented in this paper was to develop a bone segmentation algorithm to be used on MR images for CAS that would be quicker and less user dependent than manual segmentation. The accuracy of the algorithm was tested using registration experiments and compared to manually segmented MRI data and CT thresholded data.

2. Materials and methods

2.1. MR imaging

Imaging was carried out on a 1.5 T Magnetom Vision scanner (Siemens, Erlangen, Germany). A 3D double echo FISP sequence was used. All images were acquired using a six-element phased-array spine coil with three consecutive elements active. The imaging parameters were as follows: echo times were 2.5 ms and 4.9 ms, repetition time was 7.3 ms, voxel size was $1 \times 1 \times 1\ \mathrm{mm}^3$ with a matrix size of $256 \times 192 \times 96$. The slice direction was sagittal, with the frequency direction, head to foot and phase, anterior to posterior. A selective radiofrequency excitation pulse with flip angle 5° was used for all studies. The pixel bandwidth was set at 651 Hz and the number of acquisitions was four. A regional saturation band was placed across the anterior of the body to reduce motion artefacts from breathing. Images were filtered for noise using a wavelet filtering technique (Donoho 1995) and corrected for non-uniform signal intensities using the low-pass filtering technique of Tofts *et al* (1994).

The tissue contrast from the sequence was T_1 weighted with the timing of the two echoes coinciding with the lipid and water signals precessing out of phase and in phase, respectively. This phase difference is due to a chemical shift of 3.5 ppm between the Larmor frequencies of the hydrogen atoms in water and lipid. The contrast resulting from these different echo times was as follows.

There was good contrast between the vertebral bodies and inter-vertebral discs in the images from the first echo (figure 1(a)). However the edges of the posterior structures were not well defined in these images.

The images from the second echo showed clear definition of the posterior structures (figure 1(b)). However there was less contrast between the vertebral bodies and inter-vertebral discs.

To achieve good contrast between all the bone and soft tissue, the two images were combined. This method is described briefly here, for more details see Hoad *et al* (2001). The images were combined using the statistical variance of a $3 \times 3 \times 3$ region for each pixel of the first-echo data. Any pixels with high variance signal above a threshold level, V_1, were replaced (in the original data) with the matching pixels from the second-echo data. The high variance data appeared mainly around the edges of the posterior structures (figure 1(c)). By substituting these data for the corresponding data from the second echo, which were of relatively constant signal intensity, the definition of the posterior structures was improved. The threshold level, V_1, was set interactively by the user. The level was set such that the posterior surfaces of the bones of the spine were clearly defined, whilst maintaining relatively constant signal intensity in the bone marrow (see figure 1(d)).

2.2. Segmentation algorithm

The MR images after processing have low signal intensity in the bone and relatively high signal intensity in the surrounding soft tissue. The problems with segmenting these data are numerous; the anterior edge of the vertebral body is very poorly defined because of poor

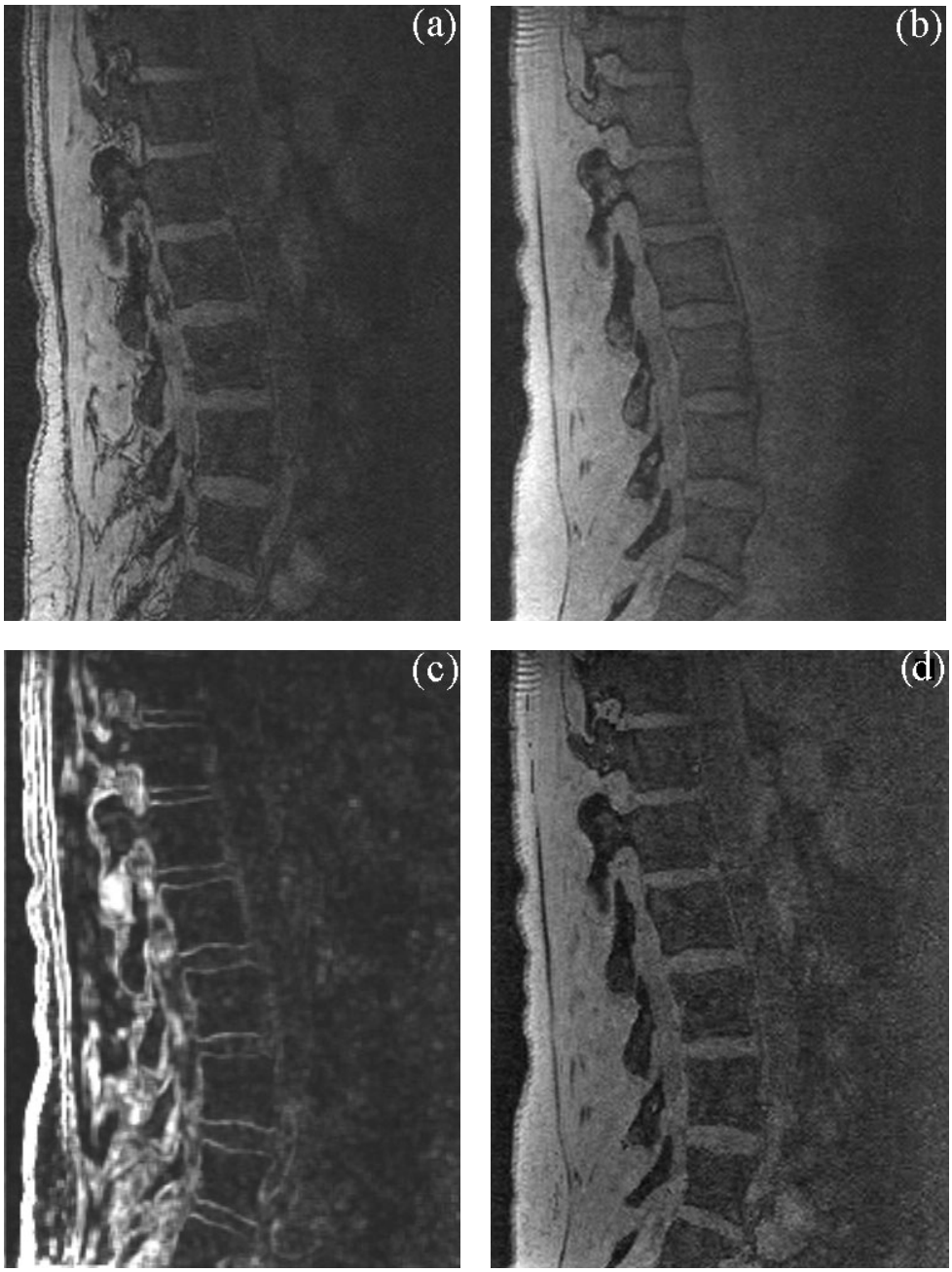

Figure 1. Examples of a typical sagittal slice MR image from the FISP sequence including processed data. (a) echo 1, (b) echo 2, (c) statistical variance data for slice and (d) resulting combined data showing good overall contrast between bone (dark) and surrounding soft tissue (light).

SNR in that region, bone signal intensity is not uniform in the thin spinous process regions and the shape of the bone changes rapidly from slice to slice. This means that very simple segmentation techniques of thesholding or edge detection do not work well on these data.

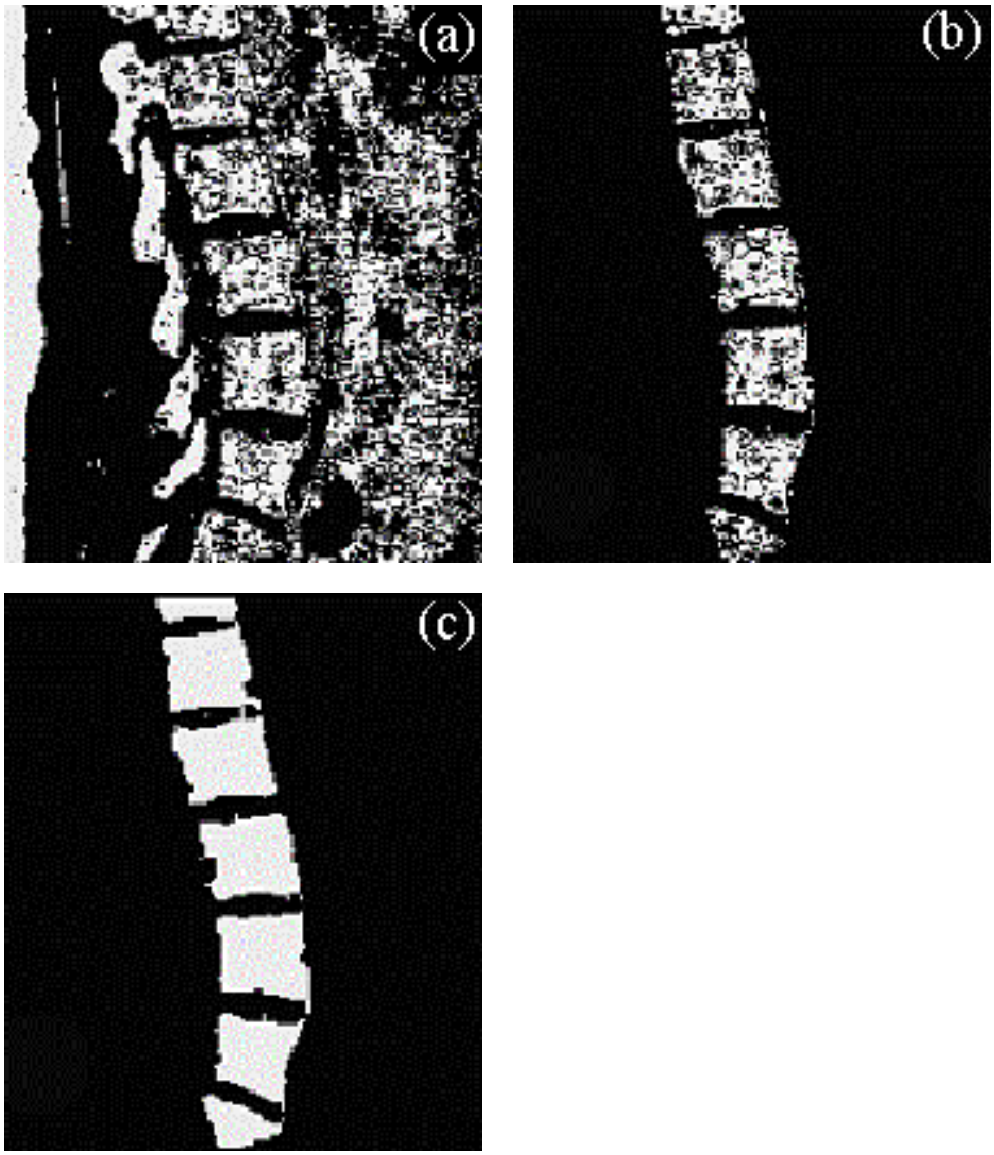

Figure 2. Single sagittal slice through data showing various stages of the algorithm for segmentation of vertebral bodies. (a) Mask of thresholded data, (b) data after elliptical column defined and (c) segmented vertebral bodies after holes filled using morphological filtering.

We have developed a segmentation algorithm that combines thresholded region growing with morphological filtering and masking using set shapes. It can be split into three sections; segmentation of the vertebral bodies, segmentation of the posterior structures and manual corrections.

2.2.1. Segmentation of the vertebral bodies. The contrast between the anterior surfaces of the vertebrae and the surrounding soft tissue is poor, making it impossible to separate the bone and soft tissue using a simple threshold (figure 2(a)). Since the anterior surface is not used

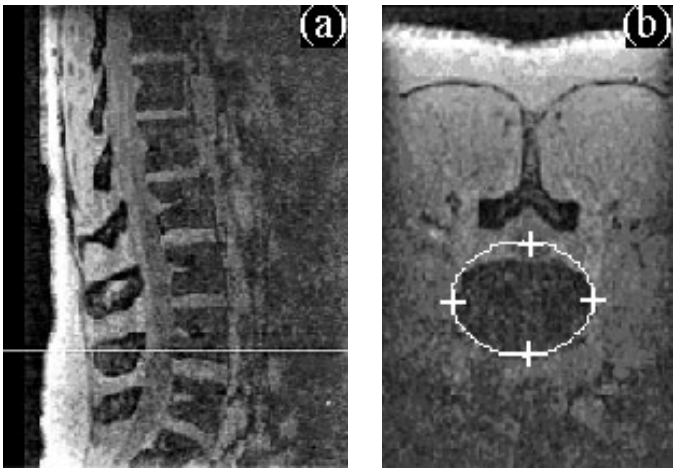

Figure 3. (a) Sagittal image showing axial slice position and (b) axial images showing the four positions drawn manually on the image (crosses) and the subsequently calculated ellipse. The ellipse shape must closely follow the anterior edge of the vertebral body for a good segmentation outcome; however, the fit to the posterior edge is not important.

for surface matching during surgery, it is possible to approximate this boundary. 1–2 ellipses were defined for each vertebral level in the axial plane. These ellipses were defined by four manually selected points and were aligned with the anterior edge of the body as closely as possible. The posterior edge of the vertebra was not matched carefully but was fully included in the ellipse. An example is given in figure 3. Interpolation was then used to produce a stack of ellipses enclosing a 3D volume, M_1, which included the entire vertebral body and disc area of the spine. A threshold T_1 was then applied to all the pixels within the mask M_1 to create a new mask M_2, i.e.

$$M_2(x, y, z) = \begin{cases} 1 \cdot M_1(x, y, z) & S(x, y, z) \leqslant T_1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where S is the signal intensity from the processed images (figure 2(b)). The threshold level T_1 was set such that most of the bone was included (especially in the posterior structures), however the amount of noise included was minimized. A morphological closing filter with a cubic $3 \times 3 \times 3$ kernel was applied to M_2 in order to fill any holes (figure 2(c)).

2.2.2. Segmentation of the posterior structures. An initial mask M_3 was generated by removing the vertebral body region from a thresholded dataset, i.e.

$$M_3(x, y, z) = \begin{cases} 1 \cdot (1 - M_1(x, y, z)) & S(x, y, z) \leqslant T_1 \\ 0 & \text{otherwise} \end{cases}. \tag{2}$$

All positive pixel values in M_3, lying between a third of the distance from the posterior edge of the ellipse in the axial plane and the anterior edge of the image region were set to zero. The resulting region M_3 is shown in figure 4(a). To generate automatic seed points for region growing, a morphological open filter with a cubic $5 \times 5 \times 5$ kernel was applied to M_3. Application of the filter resulted in seed points close to the facet joints of the posterior structures and also in the noise layer outside the body (figure 4(b)). Seed points lying within the noise region were automatically removed (figure 4(c)). Seeded region growing was then carried out for all the remaining seed points and this defined the posterior structures in the

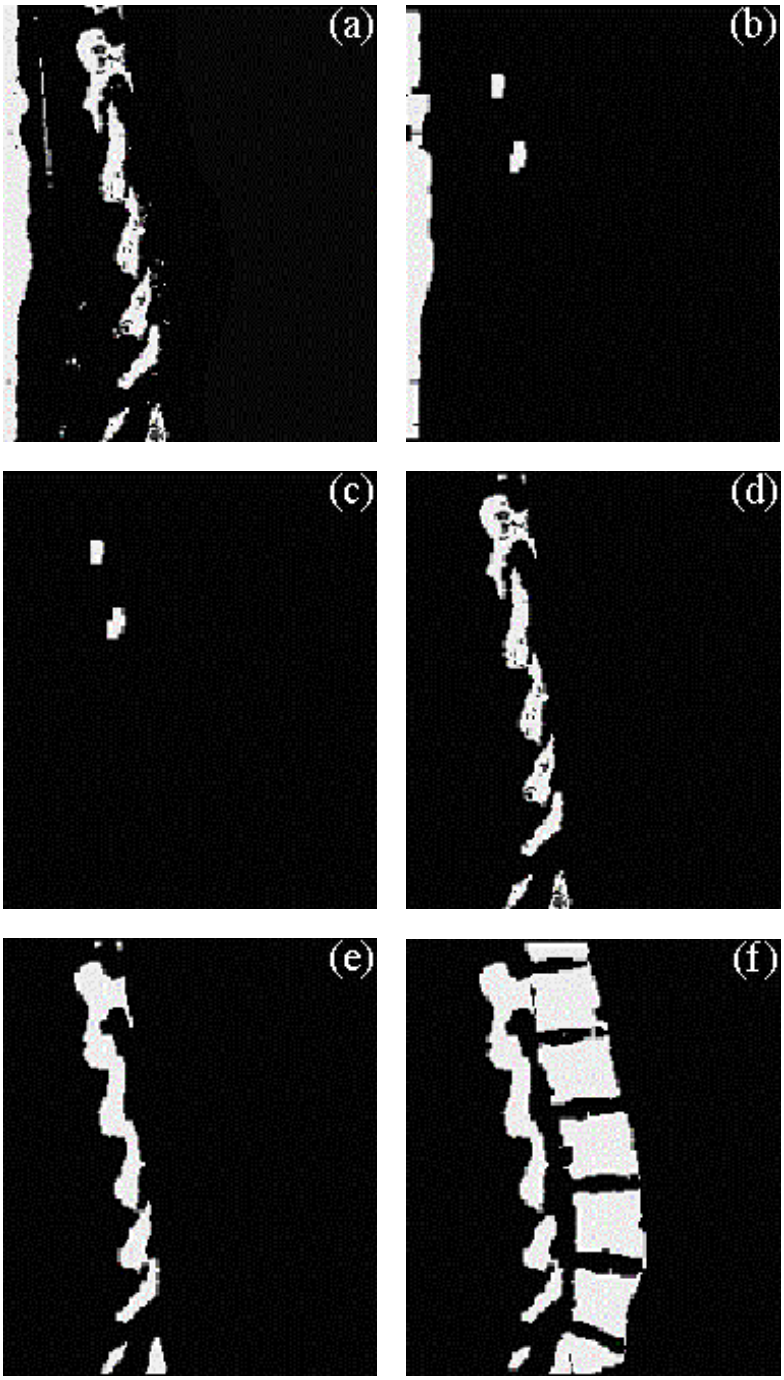

Figure 4. Single sagittal slice through data showing various stages of the algorithm for segmentation of posterior structures. (a) Thresholded data with vertebral body data removed. (b) Seed point generation using morphological open filter. (c) Seed points left after removal of noise region. (d) Data after region growing from seed points. (e) Data after holes filled by morphological closing filter. (f) Segmented data from vertebral bodies and posterior structures together.

mask M_3 (figure 4(d)). A morphological closing filter, with cubic $3 \times 3 \times 3$ kernel was used to remove any holes (figure 4(e)). Finally, masks M_2 and M_3 were added together to create the final segmented dataset, M_4. Any holes between datasets M_2 and M_3 were filled in M_4 (figure 4(f)).

2.2.3. Manual corrections. Some preliminary experiments, carried out on healthy volunteers, indicated that the thin spinous process region of the vertebra was not always well defined by the segmentation algorithm. The tip of this structure was missed sometimes and a final, manual, correction to this was needed to complete the segmentation. This manual correction was only applied to a few slices for each vertebra, when the main algorithm failed to completely define the spinous process region.

2.3. Assessment of segmentation accuracy

The aim of segmenting the images is to obtain an accurate 3D model of the spine surface which can then be used to register image space to physical space during surgery. Any method used to measure segmentation accuracy should therefore reflect how well this task is carried out. For this reason we have chosen to assess how well the segmented MR images can be registered to a physical phantom using a commercially available CAS system (Medivision, Stratec Medical Limited, Welwyn Garden City, UK). This system uses an Optotrak system (Northern Digital,Canada) and is described in more detail elsewhere (Nolte *et al* 1995). We have used the surface-based registration algorithm (Bachler *et al* 2001) provided by the system software. A physical model of the spine was constructed from CT data. A patient already scheduled for CAS of the spine was scanned using both CT and MRI on the same day. The thresholded CT scan was then used to generate a 3D computer model of the spine. This was used to build a physical model using the rapid prototyping technique of stereolithography. The model was made out of a photosensitive resin. The physical model covered the regions L4 to S1 and was used as a phantom for the following experiments.

2.3.1. Visual assessment of registration. During surgery the quality of the registration is assessed visually. As the instrument is moved around the surgical field the computer displays three orthogonal images centred on the tip of the instrument. The registration is verified by checking that when the tip of the instrument is touching the surface, the centre of the three image planes also corresponds to a bony surface. Twelve points were marked on the physical model to use as verification points (see figure 5). Points were assessed as 'good' if the cross-hairs marking the centre of the image were touching bone and 'bad' if they were either outside or inside the bone. A poor registration was classed as one that had four or more 'bad' points out of the 12 checked. A set of 30 experiments was carried out with 25 different surface points digitized for each experiment. In each experiment, the surface registration was carried out for the CT data, the MRI data segmented manually using a live wire technique (Barrett *et al* 1997) and the MRI data segmented using the technique described in this paper. Verification was carried out after each registration and the number of poor points for each segmentation type was recorded.

2.3.2. Comparison with CT. The second method for determining registration accuracy assumes that the CT registration can be used as a gold standard. Both the segmented MRI data and the segmented CT data were registered to the phantom using the same set of digitized points. CTF_{MRI} and CTF_{CT} are the CTFs mapping the MRI and CT data into physical

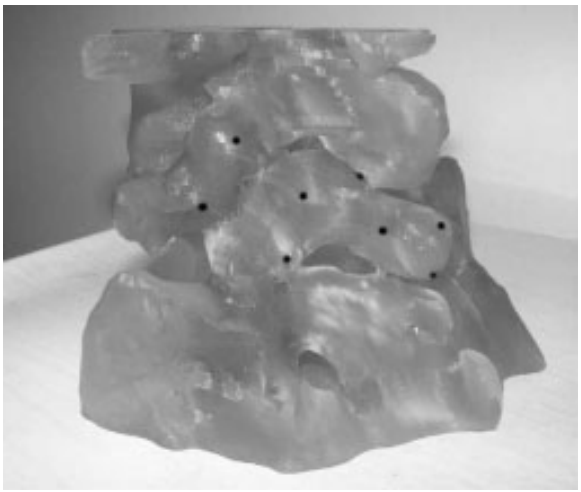

Figure 5. Picture of the physical spine model showing verification points (black dots) on L5. The model covers regions L4 to S1.

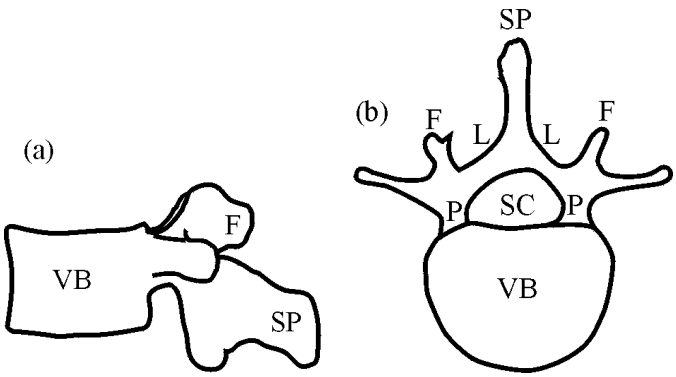

Figure 6. Different regions of the spine used in the registration experiments. (a) Sagittal view and (b) axial view. Regions are: VB—vertebral body, F—facet joint, SP—spinous process, L—laminar region, SC—spinal canal and P—pedicle.

space, respectively. The misalignment between the CT surface and the MRI surface was then quantified. Calculating the distance from a point on one surface to another surface is time consuming and in order to reduce processing time a distance map was generated from the segmented CT data by calculating the minimum distance between each pixel and the nearest surface. This distance map was defined in CT image space and had to be generated only once. In order to measure the misalignment between the registered CT and MR images, the segmented MR image was transformed into CT image space by applying CTF_{MRI} followed by $(CTF_{CT})^{-1}$. The misalignment was then quantified by randomly selecting points on the transformed MRI surface and finding their average distance away from the CT surface using the distance map. Seven regions over the surface of L5 were considered: spinous process, laminar region, facet joints, pedicles, spinal canal, vertebral body and total surface (see figure 6). Each small region had 100 points randomly selected. The total surface had 500 points randomly selected.

In order to measure the variability of the CT registration, the CTF_{CT} that represented the best surface registration match from the entire set of CT experiments was selected (CTF^{*}_{CT}). The CT surface was then transformed by CTF_{CT} followed by $(CTF^{*}_{CT})^{-1}$, and the misalignment between the transformed surface and the original surface was calculated using the distance map.

2.3.3. Effect of preprocessing and selection of threshold levels. To test the robustness of the segmentation algorithm, a further verification experiment was carried out. The threshold levels for variance, V_1 (statistical variance of $3 \times 3 \times 3$ regions in FISP first-echo data), and image segmentation threshold level, T_1, were varied between what was considered the minimum and maximum possible values that were visually acceptable. Five values uniformly spaced between these end points were selected. The segmentation algorithm was then applied and a mask generated for all combinations of V_1 and T_1. The thin spinous process region was manually segmented (once) and these data were added to the masks generated from the algorithm. The 12 points defined in the initial verification experiment were then used again to look at how well image registration had been calculated from an initial paired-point match and subsequent 25 point surface registration. Ten different sets of surface points were used to register each mask generated to the physical model. Registration accuracy was assessed by the verification method described above. No comparison was made to the CT or MR manually segmented data.

3. Results

3.1. Algorithm testing

Figure 7 shows some examples of segmented data for various healthy volunteers. The segmentation is shown as a contour overlay on the processed data. The images show that the segmentation algorithm works well on most areas of the bone. The facet joints and laminar regions are segmented accurately. These regions are particularly important because surface matching is only carried out on these two regions and the spinous process region. The spinous process region is not segmented as well. This is predominately because it is a very thin structure and partial volume effects make the edges unclear. After the image combination algorithm used to process the data, the bone marrow in the spinous process has high instead of low signal and hence is not included when the threshold level is set. As this region is small, manually segmenting this area can be carried out, as previously described, in a short amount of time.

3.2. Visual assessment of registration

The results of the registration experiments are given in table 1. Both the CT and MR semi-automatic segmented data produced 'good' registration for all experiments carried out. The manually segmented data failed to register on 10% of the experiments. On average, 10 of the 12 points were verified as good for the CT and MR semi-automatic and this figure was reduced to nine for the manually segmented MR data.

3.3. Comparison with CT

The surface distance experiments (table 2) showed that the CT experiments had smaller registration errors than the MR experiments. Both MR segmentation techniques produced errors of similar magnitude. The total surface registration error for the CT data was

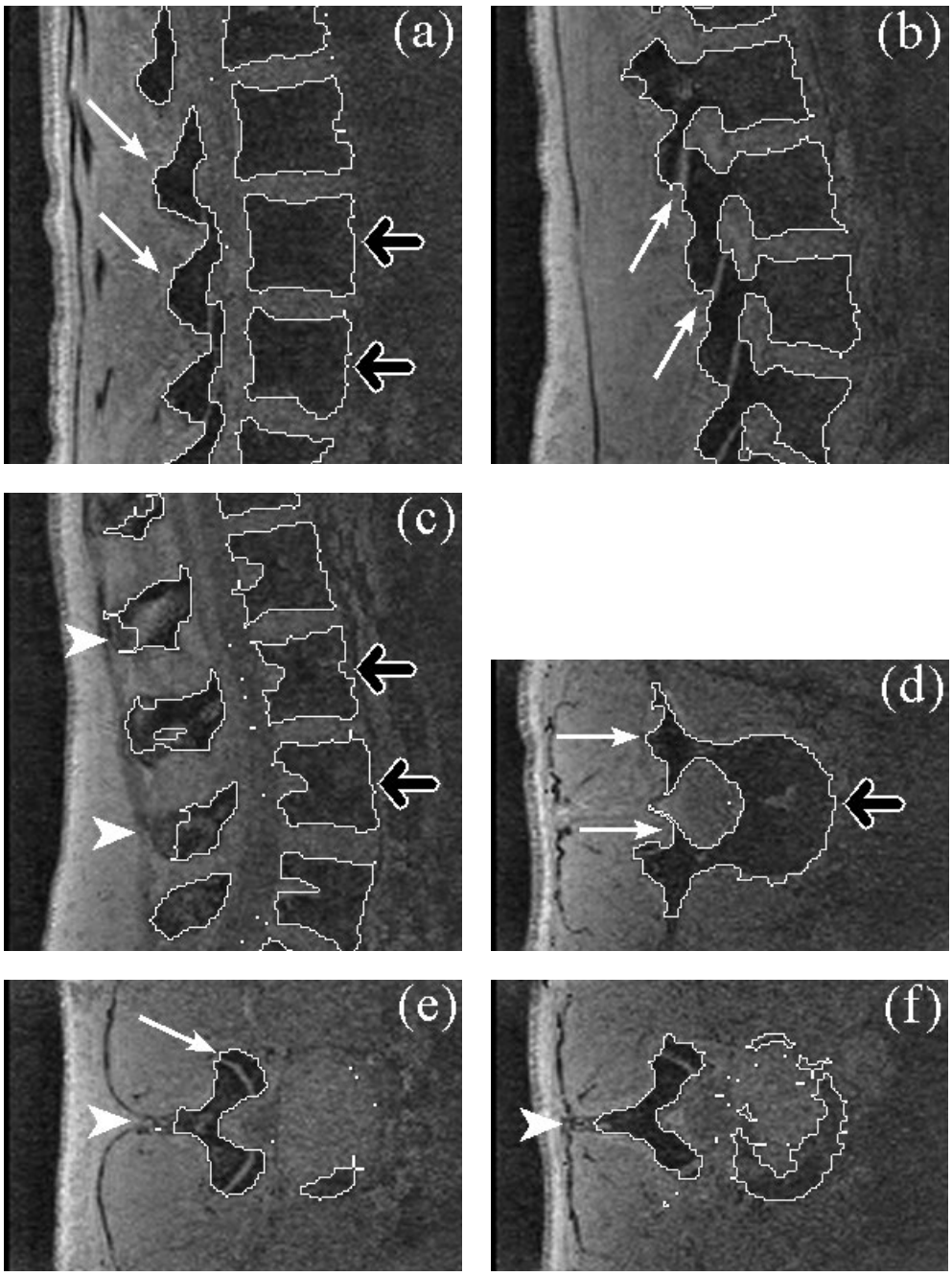

Figure 7. Images showing overlay of segmented data using algorithm for data from healthy volunteers. (a)–(c) Sagittal images and (d)–(f) axial images showing good segmentation of vertebral bodies (clear arrows), facet joints and laminar region (thin arrows). However, spinous process region is not as well defined (arrow heads).

0.66 ± 0.17 mm, for the MR semi-automatic data the error was 1.12 ± 0.15 mm and for the MR manual data 1.09 ± 0.13 mm.

Table 1. Comparison of registration accuracy results between segmentation techniques, using the verification method.

	CT	MR–manual	MR–semi-auto
Number of registrations rejected from 30 experiments	0	3	0
Average percentage of 'good' points from accepted registrations (maximum 30 experiments)	85.2%	79.4%	87.1%

Table 2. Results of the registration accuracy distance map experiment comparing segmentation techniques. All measurements are in mm. Results shown are mean ± standard deviation of mean.

Registration method	CT–CT	MR–CT (manual)	MR–CT (semi-auto)
Number of registration experiments	20	20	10
Spinous process	0.77 ± 0.36	0.89 ± 0.10	1.13 ± 0.31
Laminar	0.47 ± 0.12	0.86 ± 0.23	0.77 ± 0.12
Facet joints	0.57 ± 0.17	1.11 ± 0.16	1.10 ± 0.20
Pedicles	0.60 ± 0.22	1.18 ± 0.24	1.08 ± 0.19
Spinal canal	0.42 ± 0.14	1.27 ± 0.27	1.12 ± 0.18
Vertebral body	0.74 ± 0.29	1.25 ± 0.28	1.11 ± 0.16
Total	0.66 ± 0.17	1.12 ± 0.15	1.09 ± 0.13

Table 3. Average percentage of 'good' points from 10 registration experiments for various T_1 and V_1 levels.

	T_1-min	T_1-lower_int	T_1-centre	T_1-upper_int	T_1-max	Average
V_1-min	85	85	83	83	83	83.8
V_1-lower_int	86	83	84	83	82	83.6
V_1-centre	84	86	85	84	81	84
V_1-upper_int	82	83	79	78	80	80.4
V_1-max	76	76	78	79	76	77
Average	82.6	82.6	81.8	81.4	80.4	

3.4. Effect of preprocessing and selection of threshold levels

The experiment to test the robustness of the segmentation algorithm showed that a range of values for both T_1 and V_1 could produce good registration results (table 3). Variations between visually acceptable minimum and maximum values of T_1 and V_1 showed little variation of registration accuracy for the threshold level T_1 but a significant drop in registration accuracy for V_1 as the maximum value was reached.

4. Discussion

The segmentation algorithm described in this paper was a three-step approach and performed well at segmenting bone from surrounding soft tissue. The main algorithm (step 2) took between 5 and 10 minutes to run after the vertebral bodies (step 1) had been defined with ellipses. This meant that it was very easy to change the threshold level, T_1, for the algorithm if it was initially set incorrectly. The accuracy of the vertebral body definition does not affect

the surface registration accuracy, just the aesthetics of the 3D model presented to the surgeon on the CAS system. There was some inclusion of spinal cord tissue in the segmented data. This did not affect the registration accuracy, as this surface was not used for the registration process.

The small region of the spinous process was not segmented well on most datasets. This region is very thin and partial volume effects make the edge definition of this region unclear. Also the bone marrow in this small part of the spine has increased signal intensity compared to the rest of the spine. This is a consequence of the image combination method. Signal around the spinous process has high variance because of the rapidly changing shape of the spinous process. Therefore, low signal intensity in the bone marrow of this region cannot be achieved in the image combination. The spinous process region can be manually altered (step 3) to correct for regions missed by the semi-automatic algorithm. This is necessary for only a small number of slices (<10 normally) and can be carried out in a short amount of time. Occasionally small strips of interface tissue between the muscle and fat of the back are also included in the segmented data. This can be removed easily using manual techniques. However, if left on the model, the only inaccuracies which will occur are at the points of contact to the tip of the spinous process.

The visual verification experiment showed that there was little difference between the registration of digitized points to CT or MR generated surfaces. The MR manual segmentation surface resulted in good registration in only 90% of the experiments, whereas the CT and semi-automatic MR surfaces resulted in good registration for 100% of the experiments carried out. 100% accuracy for the CT data would be expected as the model used to digitize the point in 'physical' space was generated from this surface data. The fact that the MR semi-automatically segmented data performed as well as the CT data showed that the images and segmentation algorithm described the bone surface as accurately as the CT data.

The results from the surface distance experiment showed that there was some variation between surface alignment over different regions of L5. However as expected, one of the smallest variations in surface alignment was in the laminar region. This is because the digitized points are taken almost exclusively from this area and the algorithm for registration minimizes the distances between these points and the surface. The average total error for each modality was just larger than the in-plane pixel resolution.

The surfaces of the images were found in different ways for the CT and MRI datasets. The CT data was transformed, through 'physical space' back to CT space and the bone threshold level was set to define the 3D surface of the vertebra. For the MRI data, as the method for defining the 3D surface was not as straightforward, the mask data generated from the original segmentation process as well as the image data were transformed. This mask was displayed as a contour edge on the image data. As the mask had been truncated through the transverse processes and approximated in the vertebral bodies, the random points were only chosen from edges that appeared lying along or close to the bone surface. The larger errors associated with the MR data may be due in part to this method being more inaccurate than the CT surface method.

Larger misalignment was seen for the spinous process region of the semi-automatic method of MR segmentation than the manual segmentation method. This reflects the errors seen around this region from using the segmentation algorithm to define the bone (principally because of partial volume effects). Herring *et al* (1998) reported larger errors when using surface registration points around the spinous process for CT images and postulated that these errors were due to partial volume effects as this structure is so thin. The data in this study also show the largest registration inaccuracy around the spinous process for the CT images.

The segmentation algorithm was shown to be reasonably robust as a range of visually acceptable values of V_1 and T_1, the user dependent parameters in the algorithm, produced good registration results. Varying the parameter T_1 had almost no effect on the registration accuracy. The threshold level T_1 did not alter the matching surface significantly as changing this parameter principally altered the amount of signal included in the vertebral body region, not the posterior structures. The parameter V_1 had a much larger impact on registration accuracy, with larger values of V_1 resulting in poorer registration results. The parameter V_1 selects which parts of the processed data come from which echo. A large value of V_1 results in more data being included from the first echo around the posterior structures region. These data do not produce as clean an edge for the segmented surface as a lower value of V_1 and is almost certainly the cause for the drop in registration accuracy. It would therefore be sensible to choose a low value of V_1 for all image combinations.

As stated earlier, the physical model was generated from the CT data, therefore the fact that the surface distance experiment showed a difference of 0.66 mm in 'registered' to 'actual' surface may seem surprising. Errors from the digitization of the surface points generated from inaccurate localization of the tip of the digitization tool would have contributed to this discrepancy, by generating an inaccurate registration. The MR data were also prone to this type of registration error. Other sources of error, that could affect the registration process for the MR data, were image resolution, distortion and segmentation accuracy. The in-plane resolution for the MR images was 1×1 mm^2, for the CT data this was considerably smaller at 0.49×0.49 mm^2, allowing for a more accurate representation of the surface of the bone. The slice thickness for the CT data was 2 mm. This parameter was only 1 mm for the MR data and this may well have contributed to the good registration achieved by the MR generated surfaces, despite the poorer in-plane resolution. The registration accuracy experiments described here present a combined error from all these sources.

Ultimately, during CAS of the spine, the registration accuracy will depend on all the factors described above. The likeness of the computer generated surface of the spine to the physical spine and the accuracy of digitization of the surface points on the exposed bone will both influence the outcome of the registration algorithm. It is therefore important to collect data from surgical trials of pedicle screw placement using CAS with MR images to determine the absolute accuracy of this work.

5. Conclusions

A three-step algorithm has been described in this paper, which segments bone from soft tissue in MR images of the spine developed for CAS. Verification experiments of registering the segmented data to a physical model of a spine (generated by CT) showed the segmented MR data to be equally good for registration as 'gold standard' CT data. The segmentation algorithm was shown to be robust. Different user-defined threshold levels, varied between visually acceptable limits, had very little effect on the registration outcome. Overall, the accuracy of the registration will depend on both the similarity between actual and computer generated surfaces, and the accuracy of the digitized points used for registration.

Acknowledgments

The study was funded by the AO Foundation, Switzerland. The authors thank Rob Kerslake for his help with the MR imaging, Mike Grevitt for surgical support and Qingmao Hu for helpful discussions about morphological filters.

References

Amiot L P *et al* 1995 Computer-assisted pedicle screw fixation. A feasibility study *Spine* **20** 1208–12

Amiot L P *et al* 2000 Comparative results between conventional and computer-assisted pedicle screw installation in the thoracic, lumbar, and sacral spine *Spine* **25** 606–14

Bachler R *et al* 2001 Restricted surface matching—numerical optimization and technical evaluation *Comput. Aided Surg.* **6** 143–52

Barrett W A *et al* 1997 Interactive live-wire boundary extraction *Med. Image Anal.* **1** 331–41

Bezdek J C *et al* 1993 Review of MR image segmentation techniques using pattern-recognition *Med. Phys.* **20** 1033–48

Brandt M E *et al* 1994 Estimation of csf, white and gray-matter volumes in hydrocephalic children using fuzzy clustering of MR-images *Comput. Med. Imaging Graph.* **18** 25–34

Calmon G *et al* 2000 Automatic measurement of changes in brain volume on consecutive 3D MR images by segmentation propagation *Magn. Reson. Imaging* **18** 439–53

Choi H *et al* 1991 Partial volume tissue classification of multichannel magnetic resonance images—a mixed model *IEEE Trans. Med. Imaging* **10** 395–407

Clarke L P *et al* 1995 MRI segmentation: methods and applications *Magn. Reson. Imaging* **13** 343–68

Clarke L P *et al* 1993 MRI-stability of 3 supervised segmentation techniques *Magn. Reson. Imaging* **11** 95–106

Donoho D L 1995 De-noising by soft-thresholding *IEEE Trans. Inf. Theory* **41** 613–27

Hall L O *et al* 1992 A comparison of neural network and fuzzy clustering-techniques in segmenting magnetic-resonance images of the brain *IEEE Trans. Neural Netw.* **3** 672–82

Herring J L *et al* 1998 Surface-based registration of CT images to physical space for image-guided surgery of the spine: a sensitivity study *IEEE Trans. Med. Imaging* **17** 743–52

Hoad C L *et al* 2001 A 3D MRI sequence for computer assisted surgery of the lumbar spine *Phys. Med. Biol.* **46** N213–N220

Hoad C L *et al* 2000 Development of a 3D MRI sequence for computer assisted surgery of the Lumbar spine *ISMRM 8th Scientific Meeting and Exhibition (Denver, CO)*

Kapur T *et al* 1996 Segmentation of brain tissue from magnetic resonance images *Med. Image. Anal.* **1** 109–27

Laine T *et al* 1997 Improved accuracy of pedicle screw insertion with computer-assisted surgery. A prospective clinical trial of 30 patients *Spine* **22** 1254–8

Lemieux L *et al* 1999 Fast, accurate, and reproducible automatic segmentation of the brain in T_1-weighted volume MRI data *Magn. Reson. Med.* **42** 127–35

Liang Z 1993 Tissue classification and segmentation of MR images *IEEE Eng. Med. Biol. Mag.* **12** 81–5

Manousakas I N *et al* 1998 Split-and-merge segmentation of magnetic resonance medical images: performance evaluation and extension to three dimensions *Comput. Biomed. Res.* **31** 393–412

Martel A L *et al* 1999 Measurement of infarct volume in strokes using adaptive segmentation of diffusion weighted MR images *Lecture Notes in Computer Science* **1679** 22–31

Martel A L *et al* 1998 Assessment of 3-dimensional magnetic resonance imaging fast low angle shot images for computer assisted spinal surgery *Comput. Aided Surg.* **3** 40–4

Merloz P *et al* 1997 Computer assisted spine surgery *Clin. Orthop.* **337** 86–96

Merloz P *et al* 1998 Pedicle screw placement using image guided techniques *Clin. Orthop.* **354** 39–48

Mitchell J R *et al* 1994 Computer-assisted identification and quantification of multiple-sclerosis lesions in MR-imaging volumes in the brain *J. Magn. Reson. Imaging* **4** 197–208

Nolte L P *et al* 1995 Image-guided insertion of transpedicular screws. A laboratory set-up *Spine* **20** 497–500

Schwarzenbach O *et al* 1997 Accuracy of computer-assisted pedicle screw placement. An *in vivo* computed tomography analysis *Spine* **22** 452–8

Sonka M *et al* 1993 *Image Processing, Analysis and Machine Vision* (London: Chapman and Hall) ch 5 (Segmentation), ch 10 (Mathematical Morphology)

Tofts P S *et al* 1994 Correction of nonuniformity in images of the spine and optic-nerve from fixed receive-only surface coils at 1.5-T *J. Comput. Assist. Tomogr.* **18** 997–1003

Vannier M *et al* 1985 Multispectral analysis of magnetic resonance images *Radiology* **154** 221–4